DOCUMENT RESUME

ED 248 539 CS 208 583

AUTHOR Reed, W. Michael; And Others
TITLE Daly and Miller's Writing Apprehension Test and Hunt's T-unit Analyses: Two Measurement Precautions in Writing Research.
PUB DATE [83]
NOTE 28p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Attitudes; Higher Education; Research Methodology; *Research Problems; *Test Construction; *Test Validity; *Writing Apprehension; *Writing Evaluation: Writing Research; Writing Skills
IDENTIFIERS Syntactic Maturity; *T Units; Writing Apprehension Test (Daly and Miller)

ABSTRACT

Writing researchers must be concerned with whether the instruments they are using are constructed as soundly as possible and, if so, whether they are performing the necessary analytic procedures to interpret the results correctly. To examine the logic in the construction of J. A. Daly and M. D. Miller's Writing Apprehension Test and the incomplete analytic procedures performed in K. Hunt's T-unit analyses, two experiments were conducted with college students. Results of the first experiment showed that the inclusion of the "uncertain" response choice in the writing apprehension test can lead to faulty results, thus lessening the validity of the instrument. Findings from the second experiment indicated that the assumption of a positive, linear relationship between syntactic complexity (T-unit analyses) and writing quality is inappropriate, especially when comparing diverse groups. The two experiments suggested that (1) either writers may be misassigned to writing apprehension groups or scores may not accurately reflect their degree of writing apprehension, thus confounding results, inferences, and theories on which other studies are subsequently based; and (2) essays may be evaluated for syntactic complexity with the false assumption that a positive, linear relationship exists between syntactic complexity and quality. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED248539

Daly and Miller's Writing Apprehension Test and Hunt's T-unit Analyses: Two Measurement Precautions in Writing Research

by

W. Michael Reed
College of Human Resources and Education
West Virginia University

John K. Burton
College of Education
Virginia Tech

Nancy M. Vandett
College of Education
Virginia Tech

RUNNING HEAD: Measurement Precautions in Writing Research

W. Michael Reed
West Virginia University
HRE Microcomputer Lab
609 Allen Hall
Morgantown, WV 26506

CS 208 583

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
W. Michael Reed

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ABSTRACT

Two major concerns writing researchers need to keep in mind are, 1) are the instruments that we are using constructed as soundly as possible? and 2) if they are, are we performing the necessary analytic procedures to interpret the results accurately? In experiment 1, it was found that the inclusion of the 3-"uncertain" response choice in the Writing Apprehension Test can lead to faulty results, thus lessening the validity of the instrument. In one analysis with the 3-"uncertain" response, there was a significant F. In a subsequent run, using data modified to reflect the participants' choosing directionally (the absence of the 3-"uncertain" response choice), the F was not significant. Twelve participants changed low, medium, or high apprehension groups, indicating in the first procedure that some participants were incorrectly placed which could result in incorrect inferences and, thus, theories. In experiment 2, the often-occurring assumption of a positive, linear relationship between syntactic complexity (T-unit analyses) and quality is challenged. Findings indicate that this assumption is inappropriate, especially when comparing diverse groups. A follow-up procedure is recommended to determine the relationship between syntactic complexity and quality.

Because theories are based on results and inferences, writing researchers should always be concerned with whether the instruments that propose to measure the constructs being investigated are valid and reliable, and, if they are, that the necessary analytic procedures are performed so that the results can be adequately interpreted. If we do not have valid and reliable instruments or do not perform needed analyses, then the results we use in our inferences and interpretations may be faulty, leading us to establish theories that "can potentially affect the adequacy of [present and future] empirical work" (Daly & Hexamer, 1983, p. 157).

Two of the most widely used measurements in writing research are Daly and Miller's Writing Apprehension Test (1975) and Hunt's T-unit analyses (1970). Much research determining the effects of and on writing apprehension and the relationships between writing apprehension and other variables, such as locus of control and self-concept, has been conducted (see Daly & Wilson, 1983, for a comprehensive listing of writing apprehension studies). Likewise, in most studies that involve an analysis of a final written product, Hunt's T-unit analyses have been widely used to determine syntactic complexity. Certain concerns arise, though, because of either faulty logic in the construction of the instrument, in the case of the Writing Apprehension Test, or incomplete analytic procedures performed on the data by researchers, as in the case of T-unit analyses. Two experiments were conducted to address these concerns.

Experiment 1: Writing Apprehension Test

The potential flaw in the Writing Apprehension Test is the inclusion of the 3-"uncertain" response choice, a factor that, first, is not needed and, second, contributes to either faulty placement in writing apprehension groups (when standard deviations are used to determine low, medium, and high apprehension groups) or inaccurate total apprehension scores. Because either grouping or individual scores are used to determine effects involving writing apprehension or relationships between writing apprehension and other variables, accurate scores are crucial.

The Writing Apprehension Test (WAT), a 26-item, Likert-scale instrument, measures writing apprehension, "a construct associated with a person's tendencies to approach or avoid situations that require writing accompanied by some amount of evaluation" (Faigley, Daly, & Witte, 1981, p. 16). Writers respond to each of the 26 items by answering "1" if they strongly agree, "2" if they agree, "3" if they are uncertain, "4" if they disagree, or "5" if they strongly disagree. According to Daly and Miller (1975), each statement is labeled either "+" (if it is a high apprehension statement) or "-" (if it is a low apprehension statement). The researcher then adds up the numerical responses to the "+" statements and the numerical responses to the "-" statements. After doing so, the combined scores are inserted in the writing apprehension formula (see Note 1).

Despite the straightforwardness of the procedure, a precaution needs to be taken when using the WAT. For a Likert-scale instrument, we generally include a "no opinion" or, in the case of the WAT, "uncertain" response choice to accommodate the possibility that the respondent cannot react to the statement (Kerlinger, 1973). The "no opinion" or "uncertain" response is generally included to deal with factual information of some form. It is possible, for example, that a survey on foreign policy would include some statements that some people would not be able to respond to. The WAT, however, does not fit into the same category as these other, factual instruments. Because this instrument ostensibly measures attitudes towards writing, the statements do not relate to factual knowledge or an activity that most people have not experienced. The basic contention here is that the "uncertain" response choice is not appropriate. If the inclusion of this choice did not potentially cause problems, it might be acceptable; however, its inclusion can lead to either faulty group placement or inaccurate total apprehension scores.

The measurement problem centers on the following possibility: a respondent can answer "3" to all statements, receive a score of 78, be put in the medium apprehension group (when grouping is based on standard deviations, as suggested by Daly and Miller (1975)) or have a writing apprehension score, without having to give any indication of his or her actual degree of writing apprehension. Studies indicate that people will often choose the "uncertain," "no opinion," or middle choice.

According to Nunnally (1967), there are several response styles, one of which is called cautiousness. This style is a tendency not to answer specifically when the correct answer is in doubt. This style is reflective of personality, rather than the inability to answer. Although this style will not affect the reliability of the instrument (for these 83 participants, the WAT was reliable for males (0.95) and for females (0.93)), it will make the instrument less valid. It is common knowledge that an instrument can be reliable, but not valid; however, in any research study, the instruments must be valid (Kerlinger, 1973; Nunnally, 1967). This possibility indicates that, by including the "uncertain" response choice, we are contributing to the possibility that some people are not accurately placed in writing apprehension groups, especially if there are many occurrences of choosing the "uncertain" response.

Grouping by Apprehension Scores

Generally in writing apprehension studies, writers are grouped based on where their WAT scores fall in relation to the standard deviation, as suggested by Daly and Miller (1975). If their scores are more than one standard deviation below the mean, they are labeled "low apprehension"; if their scores are one standard deviation or within, above or below the mean, they are labeled "medium apprehension"; or if their scores are more than one standard deviation above the mean, they are labeled "high apprehension." The virtue of standard deviation is its relationship to the standard normal curve. Given a trait assumed to be

normally distributed throughout the population (e.g., intelligence, anxiety, and apprehension), 68 percent of the scores will be within one one standard deviation, above or below the mean. In essence, in writing apprehension studies, 16 percent of the writers will be labeled "low apprehension," 68 percent "medium apprehension," and 16 percent "high apprehension." Because the WAT has not been "normed," the actual scores (and, therefore, the mean and standard deviation) will vary from sample to sample. When using an analysis of variance, this grouping is necessary.

These properties of the normal curve are critical to writing apprehension studies especially if we have included the "uncertain" response choice, as Daly and Miller prescribe. Often, though, we are measuring writing apprehension in order to determine the relationship between it and other factors, such as writing skills, preparation, sex, SAT scores, or locus of control. For example, if we are attempting to determine the relationship between writing apprehension and writing skills, we would relate the writing skills scores and writing apprehension groupings. If the low apprehension group members have higher scores than the medium apprehension group members, then we can infer that there is a negative relationship between writing apprehension and writing skills: as writing skills increase, writing apprehension decreases. The major insight here is that with each respondent who is placed in one of the three writing apprehension groups go that person's scores on other instruments, which are used to

determine the relationships between writing apprehension and the constructs these other instruments are measuring. The reliability of measures, as well as the probability of detecting "true" relationships, is affected by the amount of variance. If respondents are "over-placed" in a medium writing apprehension group because of answering "uncertain" to several statements, then our confidence (in terms of reliability) in their placement and subsequent relations drawn from that placement are lessened. If this is the case, then, as Daly and Hexamer (1983) state, we are not being totally "conscious of methodological issues that can potentially affect the adequacy of empirical work" (p. 157).

When attempting to determine relationships between writing apprehension and other factors, grouping by standard deviation, as Daly and Miller prescribe (1975) is not appropriate, especially because, rather than an analysis of variance, multiple linear regression is the appropriate procedure. This procedure deals with each writing apprehension score and each score on other instruments. Grouping according to standard deviation tends to lose information, because a mean for each group is calculated and then that mean is compared to the means of other instruments. Regression, when it can be used in place of the analysis of variance, provides more accurate results.

Procedures

Although it is not often that a respondent to the WAT will answer "uncertain" to all 26 items, it is possible that a large number of

respondents will answer "uncertain" to several of the statements. In experiment 1, 42 of the 83 participants responded "uncertain" to seven or more of the 26 statements (see Table 1). In the analysis of variance, if some scores are close to the group "break-off" scores (again determined by standard deviations), then the "uncertain" responses may contribute to inaccurate placement. In the regression procedure, it obviously does not affect placement, but it does not accurately reflect a person's degree of writing apprehension. When using either procedure, relationships or effects are determined by analyzing apprehension scores and scores on other instruments.

We wanted to determine whether if we modified the scores of all the participants to reflect their choosing directionally, if they had been forced to, would the results differ statistically from the results if we allowed their "uncertain" responses to remain intact? The scores of all 83 respondents were modified so that all "uncertain" responses ("3") were changed to a "2" if their original total writing apprehension scores were below the mean or changed to "4" if their original writing apprehension scores were above the mean. We wanted to determine whether the relationships between writing apprehension and other variables changed. If so, we could then infer that including the "uncertain" response choice could conceivably result in potentially inaccurate information concerning the relationships between writing apprehension and other variables. Because regression allows us to use each participant as an experimental unit (in the case of this

experiment, 83) versus the three (low, medium, and high apprehension groups) allowed by an analysis of variance, we chose regression because it is, perhaps, a more accurate procedure.

(Insert Table 1)

Results and Discussion

In the original multiple linear regression, when writing apprehension scores were regressed on writing skills, preparation (past writing courses), and sex, there was a significant F ($p = 0.05$), with skills as the only statistically reliable predictor of writing apprehension ($p = 0.007$): as skills increased, writing apprehension decreased (see Table 2).

After the scores were modified (as explained earlier), another multiple regression was run, regressing writing apprehension scores on writing skills, preparation, and sex. In this run, however, the F ($p = 0.07$) was not statistically significant. Based on these two analytical procedures, we could make different inferences. With the scores including "uncertain" responses, we would reject the null hypothesis. With the scores reflecting directional responses, we would retain the null hypothesis. Also, as can be seen in Table 2, both the beta's for writing skills and the overall R-squared figures differ, which would again affect the inferences we might make. Following this, we decided also to look at the changes our modified scores would have if we had used an analysis of variance and standard deviation to group our

participants, giving us low, medium, and high apprehension groups. Twelve participants would have changed writing apprehension groups, indicating the possibility that the effect that certain variables have on writing apprehension may be different from the effect when the "uncertain" response-choice is included; we could attribute that difference to changes in group placement.

(Insert Table 2)

We feel our results support the need to eliminate the "uncertain" category in favor of a forced-choice format, especially when we are attempting to determine relationships between writing apprehension and other variables. We feel that this procedure will more accurately place writers in writing apprehension groups or give more accurate scores. Because writing is an activity that most people have engaged in numerous times, this instrument does not require an "uncertain" response choice. By including such a response choice, we are contributing to potentially faulty interpretations.

Experiment 2: Hunt's T-unit Analyses

In studies involving syntactic complexity measures, faulty procedural decisions by individual researchers may occur. Although T-unit analyses are measures of syntactic complexity only, there is an assumption that syntactic complexity is a major indicator of the quality of writing, thus leading some writing researchers to assume there is a positive, linear relationship between quality and syntactic complexity.

Often this assumption stands when comparing similar groups, but the assumption may be inappropriate when comparing diverse groups. There is often an inherent assumption when using T-units that syntactic complexity also reflects quality. Although Hunt, in his report, never claims that T-unit analyses reflect quality, his findings can be misinterpreted because they are listed for grade levels. An understandable link by writing researchers to equate syntactic complexity and quality is based on the positive relationship between T-units and age (as reported by Hunt) and the fact that as we get older we tend to write better. The fault with T-unit analyses does not lie with Hunt, but with writing researchers when we imply that differences in syntactic complexity also means differences in quality. This implication can keep us from performing the necessary analytic procedures to determine what part syntactic complexity plays in the quality of writing.

In experiment 2, four intact freshman English classes of basic writers and four intact freshman English classes of honors writers at a southeastern university took part in a human memory/writing research study (see Note 2). The analytic procedures involved relating certain memory variables and syntactic complexity (based on Hunt's T-unit analyses, 1970). Of the eight writing groups, the group of basic writers who had the least experimental advantage outperformed all of the other seven groups in terms of syntactic complexity (see Note 3). When comparing basic writers with honors writers, we could claim that

basic writers write more syntactically complex sentences, and thus better sentences, than honors writers. But, that, in fact, may not be the case (see Table 3).

One explanation lies with the limitation of T-unit analyses: the T-unit is perhaps so sterile a measure that it does not take into account the quality of writing. By looking simply at T-unit figures, we are not only potentially ignoring quality of writing but also the sense, sophistication, and variety of the sentences. The T-unit takes into account the number of words per T-unit (an independent clause with any related dependent clauses and embedded modifiers; Hunt, 1970), the number of clauses per T-unit, and the number of words per clause. What it does not take into account is a stronger sense of sentence variety: are the dependent clauses introductory adverbial clauses? are dependent clauses noun or adjective clauses? would an adjective clause be better if it were made into a one-word adjective preceding the noun it is modifying? are multiple sentences of the same length better than some shorter and some longer sentences? The concerns are numerous.

Another explanation relates to methodology. Hunt's results are based on age groups. His methodology and the inferences based on his results are highly dependent upon using age-groups, rather than other, more distinguishing groupings, for example, writing ability groups (see Note 4). Hunt stratified people according to ages and then randomly chose his samples so that writing ability within each group was adequately represented. When reporting mean scores, we basically

"hide" the extreme scores, although the extreme scores (the lowest and the highest) obviously affect the mean scores. This procedure is sound, and Hunt was correct in his analyses. But, this procedure is correct only when writing ability is controlled, such as grouping according to age or just comparing the same writing ability groups with each other. There is a problem, though, when we attempt to determine syntactic complexity differences of varying abilities.

In experiment 2, although our inference that basic writers write more syntactically complex sentences than honors writers is correct when we use Hunt's interpretive procedures, intuitively we know that basic writers, in fact, do not write the quality and variety of sentences that honors writers compose. In fact, the syntactic complexity of basic writers may just reflect high complexity that they are having difficulty controlling, thus resulting in low quality. It may just be "unpolished confusion." Either we can stop at this point and attest a surprising finding or we can decide to look at the data more closely. When accepting the first alternative, we are often assuming that there is a positive, linear relationship between quality of writing and syntactic complexity; that, in fact, greater syntactic complexity means better quality. If this is the case, then we must accept our finding that basic writers do, in fact, write more syntactically complex, and thus better quality, sentences than honors writers. But then, intuitively, we know this is questionable (see Table 3).

(Insert Table 3)

In this experiment, we chose the second alternative, to look at the data more closely. Because there is a discrepancy in the quality of sentences by basic writers and by honors writers, we felt we needed to look more closely at the relationship between quality and syntactic complexity for each writing ability level group. Take, for example, the following groups of sentences. Often a basic writer writes sentences similar to the following: I went to town where I bought a book and two hours later saw a movie when I went home and went to bed. A more skilled writer, such as an honors writer, might break up that sentence by writing the following sentences. When I went to town, I bought a book. Two hours later I saw a movie, after which I went home and went to bed. The second writer has obviously written better, more ideationally complex sentences than the first writer. The first writer, however, will score higher on syntactic complexity than the second writer: a mean of 24 words per T-unit versus a mean of 12.5 words per T-unit. The first writer will also score higher on clauses per T-unit (3 per T-unit versus 2 per T-unit) and words per clause (8 per clause versus 6.25 per clause). The second writer has displayed a better use of dependent clauses and introductory adverbial phrases and greater sentence variety, all of which the first writer has not.

When comparing the syntactic complexity of such diverse groups and implying a relationship between syntactic complexity and quality, we are confounding the problem by not isolating each writing level group. In fact, we found that there were not reliable differences

between quality scores and measures of syntactic complexity when performing an analysis of variance with quality scores as the independent variable and syntactic measures as the dependent variables (see Figures 1, 2, & 3). As Table 3 indicates, there is a positive relationship between quality and syntactic complexity, but only for basic writers. For honors writers the relationship is not apparent. This discrepancy supports the need to isolate groups and then perform analytic procedures to determine the relationship. A scattergram and/or analysis of variance will clarify this relationship.

(Insert Figures 1, 2, & 3)

The issue filters down to, what is it that we think syntactic complexity measures? If we are equating syntactic complexity measures with quality, then obvious precautions need to be taken. Without some follow-up procedure, the measures of syntactic complexity may, in fact, be totally invalid; we may be using a measurement that is not measuring what we want it to measure. In fact, in many studies, T-unit analyses may not be necessary if we're interested only in the quality of writing. And yet, there is another concern: if we're not using T-unit analyses to indicate quality, then why perform T-unit analyses? Simply put, is complexity worth knowing if it's not of good quality?

It is because of this possibility, and, certainly, when we are comparing two such distinct groups as basic writers and honors writers, that a quality of writing score must be included. Otherwise

by not determining the relationship between quality and syntactic complexity and considering other procedures to gain a better understanding of what we feel syntactic complexity measures, we must deal with results that are intuitively unsound or, perhaps, not worth interpreting.

Implications

As Daly and Hexamer (1983) state: "When studies are designed without sufficient concern for the adequacy of procedures, conclusions can be inherently flawed" (p. 157). Without taking into consideration the precautions raised here when using either the Writing Apprehension Test or T-unit analyses, "inherently flawed" conclusions can be made: either writers may be misassigned to writing apprehension groups or scores may not accurately reflect their degree of writing apprehension, thus confounding results, inferences, and, in the long run, theories on which other studies will be based; and essays may be evaluated for syntactic complexity with the assumption that a positive, linear relationship exists between syntactic complexity and quality. We feel that the solutions posed in these two experiments will provide more accurate results on which to base inferences.

Notes

[1] The formula that appeared in Research in the Teaching of English (1975) was incorrectly reported. The formula that appeared in this issue was

78+Positive Statements-Negative Statements.

The formula should have read

78+Negative Statements-Positive Statements.

For example, if we used the reported formula, a "high apprehensive" writer who responded "agree" to all 13 high apprehensive statements ("+" statements) and "disagree" to all 13 low apprehensive statements ("-" statements) would have a score of 44 (78+26-52=44), a score to surely place him or her in the low apprehensive group. However, if we use the same responses and apply them to the correct formula, the score is very different and much in line with the responses that were given (78+52-26=104).

[2] All groups had the same topic (an event that had had an emotional impact on them), audience (trusted adult), and mode (narrative). There were three experimental treatments (writing down ideas during the prewriting stage that was guided by an invention strategy based on Burke's pentad/ratios; writing down ideas during the prewriting stage that was guided by the writer's own mtehod of generating ideas; and not being able to write down ideas during the prewriting stage, but merely thinking about ideas that was guided by the Burke invention strategy) and one control treatment (not being able to write down ideas during the prewriting stage, but merely thinking about ideas which was guided by the writer's own usual method of generating ideas).

[3] This finding suggests the puzzling question posed from other studies investigating the relationship between syntactic complexity and quality ratings (Crowhurst, 1980; Mellon, 1969; Schmeling, 1970). These studies indicate that there is not a positive, linear relationship between syntactic complexity and quality. Unlike this study, however, they do not pose analytic solutions to ensure that researchers do not disregard this faulty assumption.

[4] On p. 12 in his report, Hunt (1970) explains his procedures in selecting the sample that "would represent something close to a normal distribution of academic ability" (p. 12).

REFERENCES

Crowhurst, M. (1980). Syntactic complexity and teachers' quality ratings of narrations and arguments. Research in the Teaching of English, 14(3), 223-231.

Daly, J. A. & Hexamer, A. (1983). Statistical power in research in English education. Research in the Teaching of English, 17(2), 157-164.

Daly, J. A. & Miller, M. D. (1975). The empirical development of an instrument to measure writing apprehension. Research in the Teaching of English, 9(4), 272-289.

Daly, J. A. & Wilson, D. A. (1983). Writing apprehension, self-esteem, and personality. Research in the Teaching of English, 17(4), 327-342.

Faigley, L., Daly, J. A., & Witte, S. P. (1981). The effects of writing apprehension on writing performance and competence. Journal of Educational Research, 75, 16-21.

Hunt, K. (1970). Syntactic maturity in school children and adults. Monographs of the Society for Research in Child Development, 35, (serial no. 134).

Kerlinger, F. N. (1973). Foundations of behavioral research. New York: Holt, Rinehart and Winston, Inc.

Mellon, J. C. (1969). Transformational sentence-combining. Research Report No. 10. Champaign, IL: National Council of Teachers of English.

Nunnally, J. C. (1967). Psychometric theory. New York: McGraw-Hill.

Schmeling, H. H. (1970). A study of the relationship between certain syntactic features and overall quality of college freshman writing. (Doctoral dissertation, George Peabody College for Teachers, 1969). Dissertation Abstracts International, 30, 4970A. (University Microfilms No. 70-7641).

Table 1: Original scores on Daly and Miller's Writing Apprehension Test, the number of "uncertain" responses, and modified scores.

n = 83 $\bar{X}$ = 75	Original scores	number of "uncertain"	Modified scores
1.	56	4	52
2.	80	9	89
3.	89	9	98
4.	68	7	61
5.	61	2	59
6.	92	7	99
7.	102	8	110
8.	99	5	104
9.	60	9	51
10.	88	2	90
11.	84	7	91
12.	84	10	94
13.	67	3	64
14.	71	8	63
15.	67	4	63
16.	73	11	62
17.	50	4	46
18.	99	7	106
19.	80	11	91
20.	83	8	91
21.	42	2	40
22.	76	5	81
23.	72	8	64
24.	104	2	106
25.	79	2	81
26.	90	10	100
27.	85	5	90
28.	64	8	56
29.	94	2	92
30.	34	1	33
31.	101	5	106
32.	121	4	125
33.	64	11	53
34.	61	6	55
35.	92	7	99
36.	100	5	105
37.	74	9	65
38.	81	7	88
39.	37	0	37
40.	56	2	54
41.	74	13	61
42.	86	7	93
43.	81	9	90
44.	71	14	57
45.	49	1	48
46.	74	9	65

Table 1

n = 83 $\bar{x}$ = 75	Original scores	number of "uncertain" responses	Modified scores
47.	63	6	57
48.	58	2	56
49.	67	8	59
50.	94	4	98
51.	64	10	54
52.	85	10	95
53.	64	6	58
54.	64	7	57
55.	77	4	81
56.	51	3	48
57.	68	6	62
58.	76	5	81
59.	72	3	69
60.	56	5	51
61.	40	4	36
62.	92	2	90
63.	55	10	45
64.	93	9	102
65.	67	12	55
66.	71	11	60
67.	78	9	87
68.	73	10	63
69.	69	7	62
70.	76	11	65
71.	63	2	61
72.	103	5	108
73.	93	0	93
74.	104	3	107
75.	82	15	97
76.	84	7	91
77.	59	10	49
78.	87	7	94
79.	64	12	52
80.	83	4	87
81.	62	12	54
82.	84	9	93
83.	97	7	104

Table 2. Multiple Linear Regressions on Original and Modified Data, Means, and Standard Deviations; Writing Apprehension as Dependent Variable, and Sex, Writing Skills, and Preparation as the Independent Variables.

Original Data

Variable	b	SEb	Beta	T	p
Skills	-0.575	0.206	-0.314	-2.789	0.007
Sex	1.630	3.976	0.048	0.410	0.683
Preparation	1.558	2.766	0.064	0.563	0.575
(Constant)	90.934	10.182	---	8.931	0.000

R squared = 0.092
Standard Error = 16.509
F(3,79) = 2.683, p = 0.05

Modified Data

Variable	b	SEb	Beta	T	p
Skills	-0.727	0.269	-0.305	-2.697	0.009
Sex	2.099	5.197	0.048	0.404	0.687
Preparation	0.970	3.615	0.031	0.268	0.789
(Constant)	96.526	13.307		7.254	0.000

R squared = 0.086
Standard Error = 21.577
F(3,79) = 2.487, p = 0.07

	Original		Modified	
	Mean	Standard Deviation	Mean	Standard Devia
Writing Appre.	75.313	17.010	74.952	22.156
Skills	36.542	9.294	36.542	9.294
Sex	1.506	0.503	1.506	0.503
Preparation	1.892	0.699	1.892	0.699

Table 3. Means for several analyses of variances with quality scores, words per T-unit, words per clause, and clauses per T-unit as dependent variables; and independent variables as external storage, no external storage, strategy, and no strategy. Main effects and interactions are listed below.

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
BASIC	3.50 Quality[1]	3.33 Quality[1]	2 50 Quality[1]	3.91 Quality[1]
	13.67 Words/T[2]	12.83 Words/T[2]	11.78 Words/T[2]	17.61 Words/T[2]
	9.29 Words/C[3]	7.88 Words/C[3]	8.17 Words/C[3]	9.89 Words/C[3]
	1.47 Clause/T[4]	1.64 Clause/T[4]	1.48 Clause/T[4]	1.83 Clause/T[4]
HONORS	6.43 Quality[1]	5.96 Quality[1]	6.09 Quality[1]	5.43 Quality[1]
	15 38 Words/T[2]	14.63 Words/T[2]	13.71 Words/T[2]	13.89 Words/T[2]
	9.38 Words/C[3]	9.72 Words/C[3]	9.35 Words/C[3]	9.00 Words/C[3]
	1.68 Clause/T[4]	1.52 Clause/T[4]	1.42 Clause/T[4]	1.57 Clause/T[4]

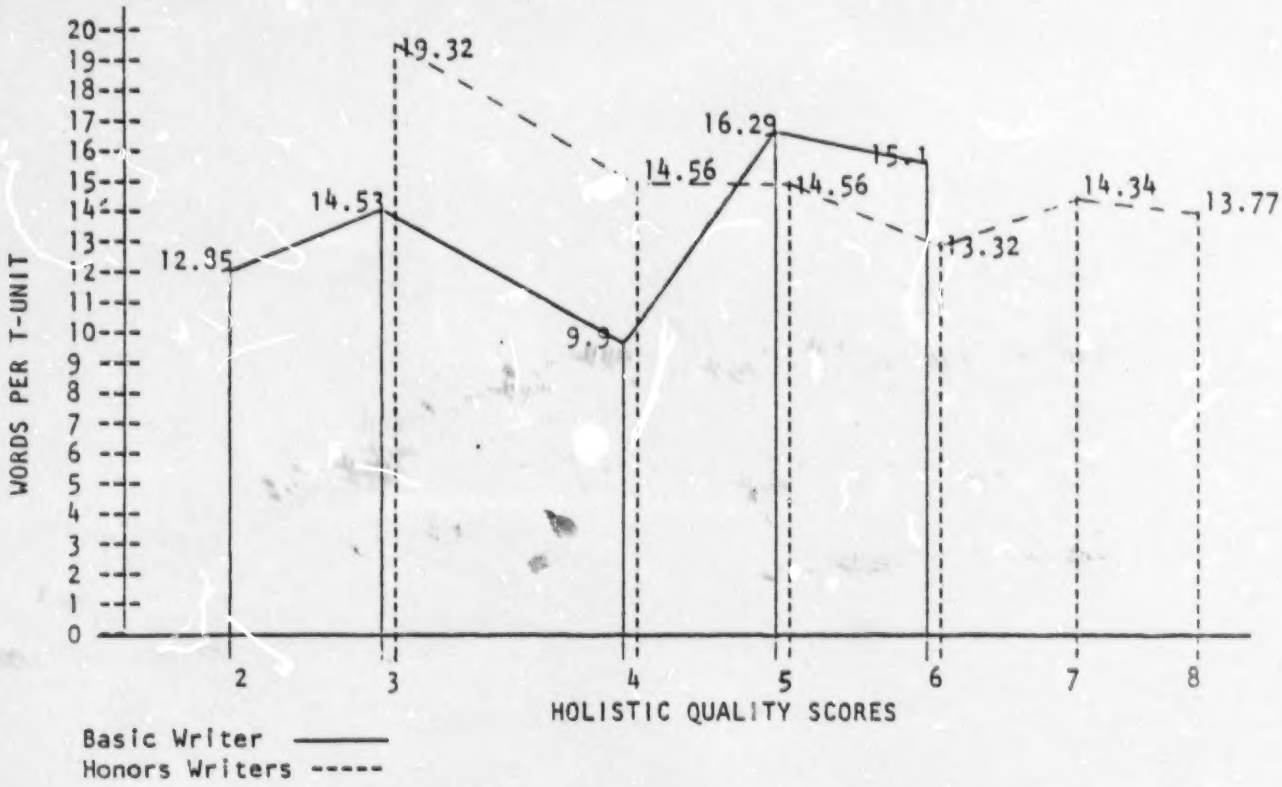
[1] $F(7,126) = 19.416$, $p = 0.000$; significant main effects ($p = 0.000$): level ($p = 0.000$), storage ($p = 0.09$), strategy ($p = 0.536$); significant two-way interactions ($p = 0.064$): level by storage ($p = 0.726$), level by strategy ($p = 0.013$), storage by strategy ($p = 0.414$); significant three-way interaction ($p = 0.081$).

[2] $F(7,126) = 2.609$, $p = 0.015$; no significant main effects; significant two-way interactions ($p = 0.009$): level by storage ($p = 0.085$), level by strategy ($p = 0.038$), storage by strategy ($p = 0.041$); significant three-way interaction ($p = 0.048$).

[3] $F(7,126) = 1.451$, $p = 0.191$; no significant main effects; no significant two-way interactions; significant three-way interaction ($p = 0.015$).

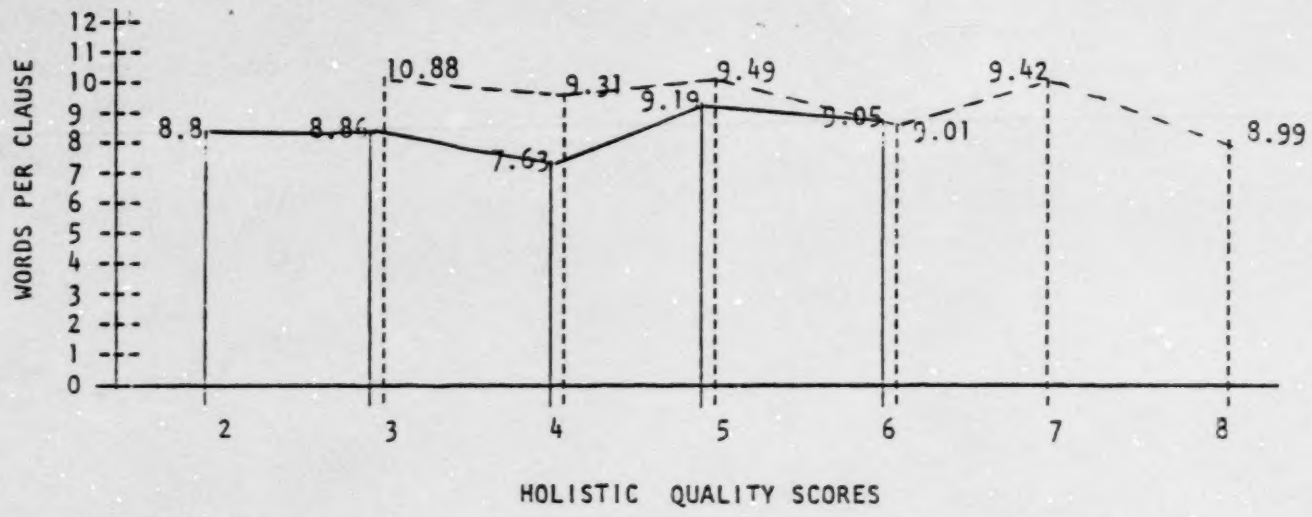
[4] $F(7,126) = 1.477$, $p = 0.181$; no significant main effects; significant two-way interactions ($p = 0.048$): level by storage ($p = 0.182$), level by strategy ($p = 0.091$), storage by strategy (p 0.076); no significant three-way interactions.

Figure 1. Graph Indicating Relationship Between Words Per T-unit and Holistic Quality Scores for Basic and Honors Writers.



NOTE: There was a significant relationship between quality and words per T-unit for basic writers ($F(4,39) = 2.2$, $p = 0.09$), but not for honors writers ($F(5,84) = 1.501$, $p = 0.2$).

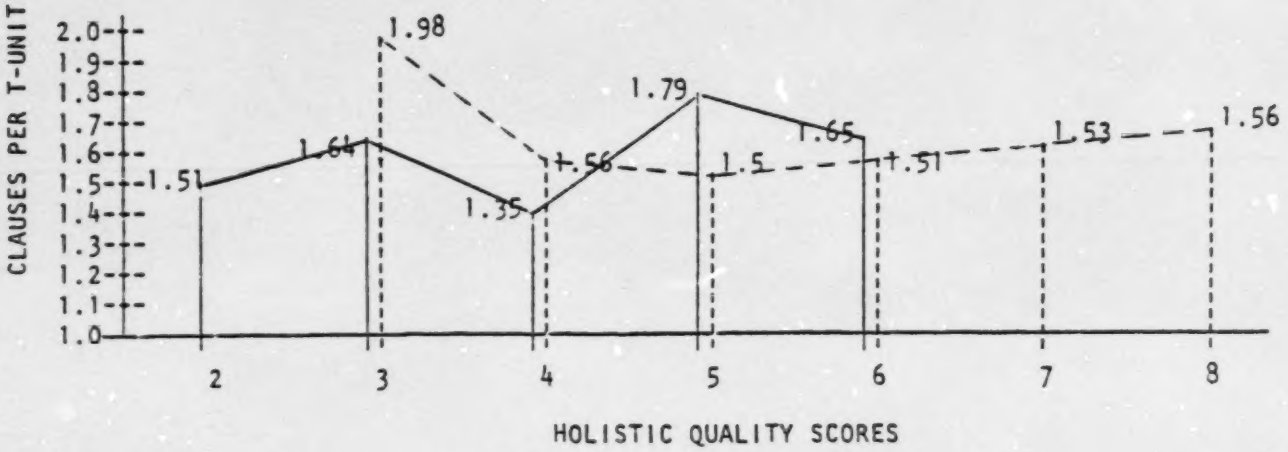
Figure 2. Graph IndicatingRelationship Between Words Per Clause and Holistic Quality Scores for Basic and Honors Writers.



Basic Writers ———
Honors writers - - - - -

NOTE: There were no significant relationships between quality and words per clause for basic writers ($F(4,39) = 0.511$, $p = 0.73$) and for honors writers ($F(5,84) = 0.515$, $p = 0.76$).

Figure 3. Graph Indicating Relationship Between Clauses Per T-unit and Holistic Quality Scores for Basic and Honors Writers.



NOTE: There were no significant relationships between quality and clauses per T-unit for basic writers ($F(4,39) = 1.069$, $p = 0.39$) and for honors writers ($F(5,84) = 0.841$, $p = 0.841$).